Title: The availability of task-specific feedback does not affect 20 km time trial cycling performance or test-retest reliability in trained cyclists

David N.Borg, John O.Osborne, Ian B.Stewart, Joseph T.Costello, Jonathon Headrick, Benjamin S.McMaster, Samantha J.Borg, Geoffrey M.Minett

Running head: Feedback and cycling performance

Abstract word count: 250

Word count: 2,750

Figures: 1

Tables: 2

ABSTRACT

Objective: This study examined the influence of the availability of task-specific feedback on 20 km time trial (20TT) cycling performance and test-retest reliability. **Design:** Thirty trained, club-level cyclists completed two 20TT's on different days, with (feedback, FB) or without (no-feedback, NFB) task-specific feedback (i.e., power output, cadence, gear and heart rate [HR]). Elapsed distance was provided in both conditions. **Methods:** During trials, ergometer variables and HR were continuously recorded, and a rating of perceived exertion (RPE) was collected every 2 km. Data were analysed using linear mixed-effects models in a Bayesian framework, and Cohen's *d* was calculated for standardised differences. The reliability of finish time and mean power output (PO) was determined via multiple indices, including intraclass correlations (ICC). **Results:** Performance, pacing behaviour, and RPE were not statistically different between conditions. The posterior mean difference [95% credible interval] between TT1 and TT2 for FB and NFB was 10 seconds [-5, 25] and -2 seconds [-17, 14], respectively. In TT2, HR was statistically higher (~8 $b \cdot min^{-1}$) in FB compared to NFB after 13 km (d = 2.08–2.25). However, this result was explained by differences in maximal HR. Finish time (FB: ICC = 0.99; NFB: ICC = 0.99) and mean power output (FB: ICC = 0.99; NFB: ICC = 0.99) in each condition were substantially reliable. **Conclusion:** The availability of task-specific information did not affect 20TT performance or reliability. Except for elapsed distance, task-specific feedback should be withheld from trained cyclists when evaluating interventions that may affect performance, to prevent participants from recalling previous performance settings.

Keywords: Exercise, pacing, reproducibility, variation, behaviour

PRACTICAL IMPLICATIONS

- The availability of task-specific feedback does not affect laboratory-based 20 km time trial performance or test-retest reliability in trained cyclists.
- Except for elapsed distance, task-specific feedback should be withheld from trained cyclists when evaluating interventions that may affect performance.
- Studies should report the type of task-specific information made available to participants during time trial performances.

INTRODUCTION

Laboratory cycling time trial (TT) tests are commonly used to evaluate interventions (e.g., ergogenic aids) and assess endurance performance.[1,2] In such research, an effective test must demonstrate high reliability in the studied population.[1,2] The test-retest reliability of TT performance in trained cyclists[3,4] has been quantified over a range of distances.[5-7] Collectively, these studies indicate that trained cyclists demonstrate highly reproducible finish times and mean power outputs (intraclass correlation [ICC] = 0.83–0.99; coefficient of variation [CV] = 0.9–2.3%).[5,8,9] Most studies describing cycling TT reliability prohibit the use of music, withhold encouragement, and standardise any feedback made available to participants (e.g., power output [PO], cadence, heart rate [HR]).[5,8,9] However, the provision of task-specific feedback between these studies is inconsistent.[2] Some provide elapsed distance,[6,10,11] while others provide real-time PO, cadence, and HR feedback, in addition to elapsed distance.[5,8,9] Further, and unfortunately, a number of studies do not report the sources of feedback made available to participants, a potentially pivotal aspect of the experimental design.[2]

In cycling TT's, task-specific feedback has been shown to influence pacing[12], oxygen consumption[13,14], and performance.[15-17] Therefore, the availability of feedback has implications for the evaluation of interventions (e.g., ergogenic aids, training programmes). It is clear that endpoint knowledge should be provided.[14,18] However, the provision of additional information—such as PO, cadence and gear—must be considered in the context of the studied population[19] and experimental manipulation.[2,20] In trained cyclists,[3,4] the availability of task-specific feedback may allow participants to replicate the cadence-gear combination of their last performance. This could produce a similar result to the previous TT effort, and lead to a null finding—independent of the intervention. As such, withholding all feedback has been recommended, to prevent any such comparison between TT attempts.[2] Alternatively, the absence of feedback that may be regularly used by cyclists during training could affect their perception of control in the task, which may also influence pacing behaviour, performance, and therefore, test-retest reliability.[12]

Determining the influence that task-specific information has on the reliability of TT performance in trained, club-level cyclists has important implications for research consistency, particularly when comparing interventions. This study aimed to examine the influence of providing or withholding feedback (i.e., PO, speed, cadence, gear and HR), in the presence of endpoint knowledge, on 20 km TT (20TT) cycling performance and test-retest reliability. It was hypothesised that the availability of feedback would result in greater reproducibility of 20TT performance compared to when no feedback was made available.

METHODS

A convenience sample of 30 (n = 8 females) participants provided written informed consent to participate in this study (Table 1). Participants were all club-level cyclists, and were considered 'trained' according to established criteria.[3,4] All experimental procedures adhered to the standards set by the latest revision of the declaration of Helsinki, except for pre-trial registration. The study was approved by the University Human Research Ethics Committee of The Queensland University of Technology (UHREC #1700001019).

A between-group design was utilised, with 15 participants in each group (Table 1). The between-group design allowed for test-retest reliability to be calculated for each condition while requiring only three participant visits—this would not have been possible using a cross-over design. Participants visited the laboratory on three occasions. The first visit involved an incremental cycling test and a familiarisation 20TT. Visits two and three involved a 20TT. During all 20TT's, including the familiarisation, participants (n = 15) in the feedback condition (FB) received elapsed distance and real-time cycling (i.e., PO, speed, cadence, gear) and heart rate (HR) information. In the other condition, no-feedback (NFB), participants (n = 15) received elapsed distance only. Males were block randomised to a condition (block sizes: 2, 2, 2, 2, 2, 2, 2, 2, 3, 3 order: alternating FB, NFB) and females were allocated 1:1, ensuring that the conditions were balanced for gender. Elapsed time was withheld for both conditions, with participants not informed of their finish times until after study completion.[2]

The time between visits was consistent within gender. Males completed their trials on average, 10 (±5) days apart. Females completed their trials at the same point in their menstrual cycle (i.e., within the first 7-days of a menstrual cycle).[21] During 20TT, no fan cooling was provided, and fluid consumption was prohibited.[2] Participants were asked to avoid alcohol and vigorous exercise 24 hours before testing, and to ensure their dietary intake, including fluid consumption, was similar the day before, and the morning of, each testing session. Hydration status was also consistent between experimental trials, as confirmed by a nude body mass (WB-110AZ; Tanita Corp., Japan) and a mid-stream urine specific gravity (U_{SG}; PAL-10S; Atagi Ci. Ltd, Japan) measurement (Table 2). Nude mass and U_{SG} were recorded at the start of each testing session. All testing was conducted in laboratory conditions (22 ± 1.6 °C; 56 ± 7% relative humidity).

During the first visit, participants were pre-screened (Exercise and Sports Science Australia, adult screening tool) and were familiarised to the study procedures. This included the 48 hour training recall, modified sleep diary[22], wellness scale[23], and rating of perceived exertion (RPE)[24]. The training recall, sleep diary and wellness scale were collected to ensure participants arrived in a similar state each testing day. Participants then completed an incremental cycling test (Excalibur Sport; Lode, Netherlands) to determine their peak oxygen consumption ($\dot{V}O_{2peak}$) and peak PO (PPO). After a 5 min warm-up, the test started at 1 $W \cdot kg^{-1}$ (females) or 2 $W \cdot kg^{-1}$ (males) and increased by 25 $W \cdot min^{-1}$ until voluntary failure. Calibrated open-circuit spirometry (TrueOne 2400, Parvo Medics, USA) was used during the test, and HR was recorded (Team 2; Polar Electro Oy, Finland). Values of $\dot{V}O_{2peak}$ and peak HR were taken as the highest 15 second average achieved, and PPO was calculated as described elsewhere.[4] After ~15 min, participants undertook a familiarisation 20TT under identical conditions to testing days (i.e., no encouragement, no fan cooling, and no fluid consumption or music). The incremental cycling test and familiarisation 20TT were completed on the same day to reduce the number of visits made to the laboratory by participants.

Participants arrived at the same time (±3 hours) for each testing day (visits two and three). A HR monitor fitted before participants completed the pre-cycling questionnaires. After a 5 min self-paced warm-up, participants commenced a flat-course 20TT on a calibrated Velotron Pro (RacerMate Inc., USA) using the 3D software (Version NB04.1.0.2101). The course had no simulated wind resistance or competitor avatars. Importantly, participants were given standardised instructions before each trial, to approach the 20TT as a race, and complete the distance as fast as possible. No encouragement was provided.[2] During cycling participants were able to select and alter their gear freely. Heart rate was continuously recorded, and a RPE was collected every 2 km. Ergometer data (i.e., power output, cadence) were sampled at 23 Hz and allocated to 1 km 'bins' for analysis. After cycling, nude mass was recorded. Cycling attire, including pedals, and ergometer settings were kept consistent within a participant across all tests.[2]

All analyses were performed in R (Version 3.5.0) using the RStudio environment (Version 1.1.447) and the 'rjags' and 'R2jags'[25], 'zoib'[26] and 'psych'[27] packages. The analysis comprised two parts. First, 20TT performance was compared between conditions. Second, the reliability of performance (finish time, mean PO and mean cadence) was calculated for each condition.

Linear mixed-effects models were used to determine differences in variables of interest before, during, and/or after cycling between FB and NFB (Supplemental 1). Models included condition, trial and condition x trial, or condition as fixed factors, and random intercept for each participant. PO relative to PPO, HR and RPE models included distance, trial, condition and their interactions as fixed factors, and a random intercept for each participant. Cycling models considered non-linear distance terms (i.e., $distance^{0.5}$, $distance^2$, $distance^3$) and distance x trial x condition as a fixed factor (Supplemental 1). Perceived wellness was modelled with beta regression[26], with condition, trial and condition x trial as fixed factors, and participant as a random effect variable. Vague prior distributions were utilised for the regression coefficients and variance parameters of all models (Supplemental 1). Posterior estimates are reported as the mean or mean difference (MD) and 95% credible interval (CI). Cohen's d (and 95% CI) was also calculated and interpreted as small 0.2, medium 0.5, and large 0.8.[25,28]

Multiple methods were used to determine the reliability of performance (i.e., finish time, mean PO, mean cadence), namely: the log-log ICC[29], standard error of the measurement (SEM)[30], minimum difference for a worthwhile change (MD_{WC})[30], and intra-individual CV[2] were calculated. Log-log ICC estimates, and their 95% confidence intervals, were based on a mean-rating (k = 2), absolute agreement, two-way random-effects model.[29] The lower bound of each ICC 95% confidence interval was used to quantify reliability as per the criteria: <0.10 virtually none, 0.11–0.40 slight, 0.41–0.60 fair, 0.61–0.80 moderate, and >0.80 substantial.[29] Full details of the data analysis methods are provided in Supplemental 1. Time trial data from the study are freely available at doi:10.17632/zxrdvwp6yr.3.

RESULTS

Performance was not statistically different between TT1 and TT2 for FB (MD [95% CI] = 10 s [-5, 25]) or NFB (MD [95% CI] = -2 s [-17, 14]), or between conditions (d = -0.04 to 0.28; Table 2; Supplemental 2). The percent change in performance was not statistically different between conditions (d [95% CI] = 0.38 [-0.37, 1.14]; Supplemental 2 and 3). Mean PO (d = -0.18 to 0.04) and mean cadence (d = 0.07 to 0.31) were not statistically different between conditions or trials (Table 2; Supplemental 2). Based on the lower bound of the 95% confidence interval, finish time and mean PO in both conditions, and mean cadence in FB were substantially reliable (Table 2). Mean cadence was moderately reliable in NFB (Table 2).

Pre-cycling nude mass (*d* = -0.14 to -0.01), U_{SG} (*d* = -0.22 to 0.34), 48 hour training recall (*d* = -0.45 to 0.42), wake time (*d* = -0.33 to 0.73) and sleep time (*d* = -0.18 to 0.60) were not statistically different between conditions or trials (Table 2; Supplemental 2). Baseline HR (Table 2) was not statistically different between conditions or trials (*d* = -0.55 to 0.07).

Pacing behaviour was not statistically different between trials or conditions (*d* = -0.04 to 0.64; Figure 1A; Supplemental 4). Within-condition, cadence in TT1 was statistically higher compared to TT2 for FB (*d* = 0.51 to 4.92; Figure 1B; Supplemental 4). Heart rate was statistically lower in TT1 compared to TT2 (i.e., within-condition) for FB from 1–14 km (*d* = -3.36 to -2.01; Figure 1C; Supplemental 4), and statistically higher in TT1 compared to TT2 (i.e., within-condition) for NFB from 1–16 km (*d* = 1.99 to 4.93; Figure 1C; Supplemental 4). In TT2, HR was statistically higher in FB compare to NFB from 14 km onwards (*d* = 2.08 to 2.25; Figure 1C; Supplemental 4). Within-condition, RPE was statistically lower for NFB in TT2 compared to TT1 (*d* = -5.26 to -0.17; Figure 1D; Supplemental 4). Nude body mass loss was not statistically different between conditions or trials (*d* = -0.07 to 0.61; Table 2; Supplemental 2).

DISCUSSION

This study examined the influence that the availability of task-specific feedback (i.e., PO, speed, cadence, gear and HR), in the presence of endpoint knowledge, has on 20TT performance and test-retest reliability. It was hypothesised that task-specific feedback would result in greater reproducibility of performance, compared to when no feedback was available. In contrast to the hypothesis, performance was not statistically different between conditions

(Table 2), and finish time and mean PO were similarly reliable (Table 2). In TT2, HR was statistically higher from 14 km onwards (~8 b·min^{-1}; d = 2.08 to 2.25) when task-specific feedback was available (Figure 1C). However, this result was not observed when HR responses during the 20TT were expressed relative to participants' maximal HR. Findings suggest that the availability of task-specific feedback does not affect 20TT performance or test-retest reliability.

The availability of real-time PO, speed, cadence, gear, and HR feedback was found to have little effect on pacing behaviour (Figure 1A) and performance (Table 2). These findings are consistent with previous research[18] that showed withholding feedback, including endpoint knowledge, did not alter performance, in a similar participant cohort to the current study (mean PPO = 384 W; mean $\dot{V}O_{2peak}$ ~56 mL·kg^{-1}·min^{-1}). Smits et al.[18] found that without feedback, participants did not exhibit a task end spurt, that is, an increase in work-rate in the final ~10% of the event. In contrast, both groups in the current study exhibited similar pacing behaviour, including an end spurt (Figure 1A). These contrasting findings are likely explained by the provision/absence of distance.[14] Results from the current study suggest that for trained[3,4], club-level cyclists, who regularly cycle short distances (i.e., 30–40 min, 20 km), the availability (or lack) of task-specific feedback does not affect 20TT performance.[14,18] A point of difference between the current study and Smits et al.[18] is the provision of test-retest reliability outcomes for each experimental condition (Table 2). The reliability results from the current study further confirm that cyclists are highly reliable in TT tasks with[5,8,9] and without[6,10,11] feedback.

From a research design standpoint, the provision of task-specific feedback in intervention studies requires careful consideration,[2] as do the verbal instructions given to participants before a TT task.[2,20] The provision of real-time task-specific information during cycling may result in a participant recalling all the decisions or events from their last trial and, depending on the researchers' instructions, the participant could simply aim to replicate the previous performance.[2] This would result in a null finding for the intervention, but only because participants were able to recall their cadence-gear settings, rather than the intervention failing to affect performance.[20] Therefore, although performance was equally reliable with and without feedback (Table 2), the authors recommend that task-specific feedback be withheld from participants during intervention studies, with the exception of elapsed distance. Methods sections of similar time trial research papers should also detail the sources of feedback made available to participants, to allow more accurate interpretations of findings.

An unexpected result from the current study was that participants who received feedback experienced a statistically higher HR during the second TT from 14 km onwards (Figure 1C). Initially, it was hypothesised that the higher HR response may have been the consequence of increased cognitive demand due to dual-tasking,[31] as participants were interpreting real-time information while maintaining their cycling work-rate. However, when HR responses were expressed relative to participants' maximal HR, these between-condition differences were not observed. In further support that the higher HR response was explained by participants' maximal HR, Holgado et al.[32] showed that an increased cognitive load during a 20 min self-paced cycling task had no impact on HR. Further, it is unlikely that trained cyclists[3,4], who regularly interpret task-specific feedback during normal training activities (Table 1), would find the feedback provided in the current study cognitively challenging.[19]

The primary limitation of the current study was the sample size. However, the sample size requirements for adequate stability in ICC and SEM estimates are equivocal.[29,30] The results from the current study are limited to a laboratory setting, without fanning and fluid intake, and to trained cyclists. The incremental cycling test and familiarisation 20TT were completed in the same visit. As such, the familiarisation 20TT could not be compared to TT1 and TT2, because performance in the 20TT may have been confounded by the incremental test—this could be considered a limitation. The number of days between TT1 and TT2 was different between genders. Although this was the result of ensuring female participants completed their trials during the same menstrual phase, this could be considered a limitation. An interesting avenue for future research would be to examine whether the effect of a between-trial period of four-weeks was larger than the performance variation observed in different menstrual phases. Future research should replicate the current study in a field-based setting, and explore the influence of the availability of task-specific feedback on performance across distances less familiar to participants.

CONCLUSION

Task-specific feedback (i.e., real-time PO, speed, cadence, gear and HR) did not affect 20TT performance or test-retest reliability. The provision of feedback resulted in a higher HR in the final-third of TT2 compared to when no feedback was available; however, this result was explained by differences in participants' maximal HR. Findings from the current study suggest that with the exception of elapsed distance, task-specific information should be withheld in intervention studies, to prevent participants from recalling the settings from a previous trial.

ACKNOWLEDGEMENTS

The authors would like to thank Ms Meg Doohan for her assistance with data collection, all participants for their involvement in the study, and the two reviewers for providing feedback that improved the interpretation of the findings and quality of the manuscript. This project was not financially supported.

REFERENCES

1. Hopkins WG, Hawley JA, Burke LM. Design and analysis of research on sport performance enhancement. *Med Sci Sports Exerc.* 1999;31(3):472-485.
2. Currell K, Jeukendrup AE. Validity, reliability and sensitivity of measures of sporting performance. *Sports Med.* 2008;38(4):297-316.
3. Decroix L, De Pauw K, Foster C, Meeusen R. Guidelines to classify female subject groups in sport-science research. *Int J Sports Physiol Perform*. 2016;11(2):204-213.
4. De Pauw K, Roelands B, Cheung SS, De Geus B, Rietjens G, Meeusen R. Guidelines to classify subject groups in sport-science research. *Int J Sports Physiol Perform.* 2013;8(2):111-122.
5. Borg DN, Osborne JO, Stewart IB, Costello JT, Sims JN, Minett GM. The reproducibility of 10 and 20 km time trial cycling performance in recreational cyclists, runners and team sport athletes. *J Sci Med Sport.* 2018;21(8):858-863.
6. Sparks A, Williams E, Massey H, Bridge C, Marchant D, McNaughton L. Test-retest reliability of a 16.1 km time trial in trained cyclists using the CompuTrainer ergometer. *J Sci Cycling.* 2016;5(3):35-41.
7. Smith M, Davison R, Balmer J, Bird S. Reliability of mean power recorded during indoor and outdoor self-paced 40 km cycling time-trials. *Int J Sports Med.* 2001;22(04):270-274.
8. Noreen EE, Yamamoto K, Clair K. The reliability of a simulated uphill time trial using the Velotron electronic bicycle ergometer. *Eur J Appl Physiol.* 2010;110(3):499-506.
9. Zavorsky G, Murias J, Gow J, et al. Laboratory 20-km cycle time trial reproducibility. *Int J Sports Med.* 2007;28(09):743-748.

10. Stone MR, Thomas K, Wilkinson M, Gibson ASC, Thompson KG. Consistency of perceptual and metabolic responses to a laboratory-based simulated 4,000-m cycling time trial. *Eur J Appl Physiol.* 2011;111(8):1807-1813.

11. Sporer B, McKenzie D. Reproducibility of a laboratory based 20-km time trial evaluation in competitive cyclists using the Velotron Pro ergometer. *Int J Sports Med.* 2007;28(11):940-944.

12. Gernigon C, Delloye JB. Self-efficacy, causal attribution, and track athletic performance following unexpected success or failure among elite sprinters. *Sport Psychol.* 2003;17(1):55-76.

13. Beedie C, Lane A, Wilson M. A Possible Role for Emotion and Emotion Regulation in Physiological Responses to False Performance Feedback in 10 Mile Laboratory Cycling. *Appl Psychophysiol Biofeedback.* 2012;37(4):269-277.

14. Jones HS, Williams EL, Bridge CA, et al. Physiological and psychological effects of deception on pacing strategy and performance: a review. *Sports Med.* 2013;43(12):1243-1257.

15. Noakes TD. Time to move beyond a brainless exercise physiology: the evidence for complex regulation of human exercise performance. *Appl Physiol Nutr Metab*. 2011;36(1):23-35.

16. Brick NE, MacIntyre TE, Campbell MJ. Thinking and action: a cognitive perspective on self-regulation during endurance performance. *Front Physiol.* 2016;7:159.

17. Marcora S. Counterpoint: afferent feedback from fatigued locomotor muscles is not an important determinant of endurance exercise performance. *J Appl Physiol.* 2010;108(2):454-456.

18. Smits BL, Polman RC, Otten B, Pepping G-J, Hettinga FJ. Cycling in the absence of task-related feedback: effects on pacing and performance. *Front Physiol.* 2016;7:348.

19. Boya M, Foulsham T, Hettinga F, et al. Information acquisition differences of experienced and novice time trial cyclists. *Med Sci Sports Exerc*. 2017;49:1884-1898.

20. Corbett J, White DK, Barwood MJ, et al. The effect of head-to-head competition on behavioural thermoregulation, thermophysiological strain and performance during exercise in the heat. *Sports Med*. 2018;48(5):1269-1279.

21. Charkoudian N, Stachenfeld NS. Reproductive hormone influences on thermoregulation in women. *Compr Physiol.* 2014;4(2):793-804.

22. Carney CE, Buysse DJ, Ancoli-Israel S, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep.* 2012;35(2):287-302.

23. McLean BD, Coutts AJ, Kelly V, McGuigan MR, Cormack SJ. Neuromuscular, endocrine, and perceptual fatigue responses during different length between-match microcycles in professional rugby league players. *Int J Sports Physiol Perform.* 2010;5(3):367-383.

24. Borg G. *Borg's perceived exertion and pain scales.* Human kinetics; 1998.

25. Mengersen KL, Drovandi CC, Robert CP, Pyne DB, Gore CJ. Bayesian estimation of small effects in exercise and sports science. *PLoS One.* 2016;11(4):e0147311.

26. Liu F, Kong Y. zoib: an R package for bayesian inference for beta regression and zero/one inflated beta regression. *R J.* 2015;7(2):34-51.

27. Revelle W, Revelle MW. Package ‘psych’. *The Comprehensive R Archive Network.* 2015.

28. Cohen J. A power primer. *Psychol Bull.* 1992;112(1):1155-1159.

29. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420-428.

30. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19(1):231-240.

31. Hidalgo-Muñoz AR, Béquet AJ, Astier-Juvenon M, et al. Respiration and heart rate modulation due to competing cognitive tasks while driving. *Front Hum Neurosci.* 2018;12.

32. Holgado D, Zabala M, Sanabria D. No evidence of the effect of cognitive load on self-paced cycling performance. *PloS One*. 2019;14(5):e0217825.

Table 1. Participant characteristics and training activities.

	Feedback (n = 15)	No-feedback (n = 15)
Age (years)	27.0 ± 5.3 (18.9–39.1)	28.4 ± 5.1 (22.2–39.5)
Height (cm)	178 ± 11 (160–191)	181 ± 9 (161–192)
Nude mass (kg)	73.9 ± 14.0 (47.5–92.6)	75.3 ± 10.7 (60.5–89.8)
$\dot{V}O_{2peak}$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	55.8 ± 8.2 (42.5–75.7)	58.6 ± 8.1 (50.0–75.3)
$\dot{V}O_{2peak}$ ($L \cdot min^{-1}$)	4.10 ± 1.02 (2.53–5.18)	4.42 ± 0.85 (3.07–5.72)
Absolute peak power (W)	368 ± 85 (220–475)	395 ± 68 (299–484)
Relative peak power ($W \cdot kg^{-1}$)	5.0 ± 0.8 (3.6–6.8)	5.3 ± 0.7 (4.0–7.0)
Heart rate max ($beats \cdot min^{-1}$)	191 ± 8 (177–203)	186 ± 10 (169–206)
Training sessions ($session \cdot wk^{-1}$), median (range) [a]	6 (2–9)	6 (3–10)
Training minutes ($min \cdot wk^{-1}$) [a]	509 ± 260 (120–960)	524 ± 361 (185–1650)
Performance level, median (range) [b]	3 (2–5)	3 (2–5)
Speed and distance feedback are available during training activities, n (%)	15 (100)	15 (100)
Regularly uses a heart rate monitor during training activities, n (%)	15 (100)	15 (100)
Regularly rides with a power meter in training activities, n (%)	4 (27)	5 (33)

Note. Data reported as mean ± standard deviation (range), unless otherwise stated. $\dot{V}O_{2peak}$ = Peak oxygen consumption.

[a] Training activity data calculated from the weekly average of the previous one-month.

[b] On a 1 to 5 classification scale (increments of 1), where performance level 5 indicates professional cyclists.[3,4]

Table 2. Experimental variables (pre-cycling and during cycling) and reliability outcomes for the feedback and no-feedback condition.

		Feedback (n = 15)	No-feedback (n = 15)
Experimental testing variables			
Cycling			
Finish time (min:s) [a]	TT1	33:48 [32:05, 35:31]	33:06 [31:25, 34:47]
	TT2	33:57 [32:14, 35:40]	33:04 [31:25, 34:48]
Mean power output (W)	TT1	249 [217, 280]	261 [229, 292]
	TT2	247 [215, 278]	261 [230, 293]
Mean cadence ($r \cdot min^{-1}$)	TT1	96 [91, 101]	93 [88, 98]
	TT2	95 [90, 100]	93 [88, 98]
Body mass loss (kg)	TT1	-0.8 [-1.0, -0.7]	-1.1 [-1.3, -0.9]
	TT2	-0.9 [-1.1, -0.7]	-0.9 [-1.1, -0.7]
Pre-cycling			
Nude mass (kg)	TT1	72.5 [65.8, 79.9]	74.9 [67.7, 82.5]
	TT2	72.9 [66.2, 80.4]	75.0 [67.9, 82.6]
Urine specific gravity	TT1	1.016 [1.010, 1.021]	1.015 [1.010, 1.021]
	TT2	1.012 [1.007, 1.017]	1.014 [1.009, 1.019]
Perceived wellness (5–25)	TT1	17 [16, 18]	18 [17, 19]
	TT2	17 [16, 19]	17 [16, 18]
48 hour training recall (au)	TT1	897 [440, 1320]	981 [617, 1345]
	TT2	628 [211, 1066]	861 [497, 1231]
Sleep onset time (24 h time)	TT1	21:52 [21:15, 22:28]	21:50 [21:18, 22:22]
	TT2	22:07 [21:30, 22:45]	21:47 [21:17, 22:18]
Sleep wake time (24 h time)	TT1	05:59 [05:05, 06:51]	05:02 [04:13, 05:52]
	TT2	06:50 [05:57, 07:43]	05:11 [04:26, 05:57]
Baseline heart rate ($b \cdot min^{-1}$)	TT1	70 [65, 75]	73 [67, 78]
	TT2	73 [67, 78]	68 [63, 74]
Reliability outcomes			
Finish time[a] (min:s)	ICC [b]	0.99 [0.98, 1.00]	0.99 [0.97, 1.00]
	SEM	0:21	0:19
	MD_{WC}	0:59	0:52
	CV	0.9	0.8
Mean power output (W)	ICC [b]	0.99 [0.98, 1.00]	0.99 [0.98, 1.00]
	SEM	7	6
	MD_{WC}	20	17
	CV	2.4	2.0
Mean cadence ($r \cdot min^{-1}$)	ICC [b]	0.96 [0.89, 0.99]	0.92 [0.77, 0.97]
	SEM	2	4
	MD_{WC}	7	12
	CV	2.1	3.3

Note. Experimental variables are reported as the posterior mean and 95% credible interval. There was no evidence of statistical differences between conditions or trials for all variables reported in Table 2.
CV = Intra-individual coefficient of variation; ICC = Intraclass correlation; MD_{WC} = Minimum difference for a worthwhile change; SEM = Standard error of the measurement; TT1 = Time trial 1; TT2 = Time trial 2.
[a] Finish time modelled in seconds, but shown in minutes and seconds for readability purposes.
[b] ICC estimate and 95% confidence interval based on a mean-rating, absolute agreement, two-way random-effects model.

Figure 1 legend.

Posterior mean and 95% credible interval power output (in 1 km bins) relative to participants' peak power output achieved during the incremental cycling test (A); cadence (in 1 km bins) (B); heart rate (C); and rating of perceived exertion (D) across the 20 km cycling time trial. FB = Feedback; NFB = No-feedback; TT1 = Time trial 1; TT2 = Time trial 2; * indicates statistically different between TT1 and TT2 for FB (i.e., within-condition difference); ** indicates statistically different between TT1 and TT2 for NFB (i.e., within-condition difference); # indicates statistically different between FB and NFB in TT2 (i.e., between-condition difference).